

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 1, 2001, 16:16:43 ; Search time 91.75 Seconds

(Without alignments)
219.987 Million cpi: updates/sec

Title: US-09-331-631A-5

Perfect score: 3326
Sequence: 1 00MOLETSQOMRRCVSCQDK.....SPRSTKQOQPLVSIIDFVG 625

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 88757 seqs, 32294092 residues

Total number of hits satisfying chosen parameters: 88757

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Query Length	DB ID	Description
1	1074.5	32.3	588	1	VCLB_GOSHI
2	1063.5	32.0	605	1	VCLB_GOSHI
3	900	27.1	573	1	GLB1_MAIZE
4	863.5	26.0	626	1	AH12_ARAHY
5	861	25.9	614	1	AH11_ARAHY
6	859.5	25.8	605	1	GLCA_SOYBN
7	854.5	25.7	639	1	GLCK_SOYBN
8	833	25.0	571	1	CVCA_PEA
9	794	23.9	524	1	SBP_SOYBN
10	759.5	22.8	459	1	VCLC_PEA
11	749	22.5	439	1	GLCB_SOYBN
12	713.5	21.5	463	1	VCL_VICFA
13	712.5	21.4	410	1	VCLB_PEA
14	692.5	20.8	445	1	CANA_CANEN
15	685.5	20.6	445	1	CANA_CANGL
16	563	16.9	436	1	PHSA_PHAVU
17	552	16.6	421	1	PHSR_PHAVU
18	493	14.8	386	1	CVCB_PEA
19	414	12.4	275	1	VCLB_PEA
20	216	6.5	1898	1	TRHY_HUMAN
21	166	5.9	1407	1	TRHY_RABIT
22	193	5.8	499	1	GLD2_ORYSA
23	192.5	5.8	124	1	VCL1_PEA
24	192.5	5.8	500	1	GLU5_ORYSA
25	191	5.7	499	1	GLUB_ORYSA
26	185.5	5.6	1549	1	TRHY_SHEEP
27	185	5.6	518	1	SSG2_AVEA
28	184	5.5	499	1	GLU4_ORYSA
29	184	5.5	495	1	SSG1_AVEA
30	180.5	5.4	498	1	GLU2_ORYSA
31	180.5	5.4	499	1	GLU1_ORYSA
32	180	5.4	496	1	GLUC_ORYSA
33	176.5	5.3	1023	1	GLT_DROME

34	176	5.3	544	1	INVO_AOTTR	P24708 aotus trivi
35	170.5	5.1	338	1	LEGB_PEA	P14594 pisum sativ
36	169.5	5.1	585	1	INVO_HUMAN	P07476 homo sapien
37	169	5.1	522	1	INVO_HYLLA	P17941 hylobates l
38	166.5	5.0	496	1	GLU3_ORYSA	009151 oryza sativ
39	165	5.0	708	1	GBF_DICDI	P36417 dictyostell
40	164	4.9	560	1	INVO_PANPA	P14591 pan panliscu
41	162.5	4.9	471	1	RU17_XENLA	P09406 xenopus lae
42	162	4.9	516	1	LEGB_GOSHI	P09800 gossypium h
43	160	4.8	835	1	INVO_PONPY	P14708 pongo pygma
44	157.5	4.7	493	1	INVO_SAGOE	P24712 saguinus oe
45	157.5	4.7	877	1	INCE_CHICK	P53352 gallus gall

ALIGNMENTS

RESULT	ID	VCLB_GOSHI	STANDARD:	PRT:	588 AA.
AC	P09801:				
DT	01-MAR-1989 (Rel. 10, Created)				
DT	01-MAR-1989 (Rel. 10, Last sequence update)				
DT	15-JUL-1999 (Rel. 38, Last annotation update)				
DE	VICILIN C72 PRECURSOR (ALPHA-GLOBULIN B).				
OS	Gossypium hirsutum (Upland cotton).				
OC	Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;				
OC	Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;				
OC	Malvales; Malvaceae; Gossypium.				
RN	(1)				
RA	SEQUENCE FROM N.A.				
RP	Chlan C.A., Pyle J.B., Legocki A.B., Dure L. III;				
RT	"Developmental biochemistry of cottonseed embryogenesis and				
RT	germination. XVIII. cDNA and amino acid sequences of the members of				
RT	the storage protein families."				
RL	Plant Mol. Biol. 7:475-489(1986).				
CC	-1- FUNCTION: SEED STORAGE PROTEIN.				
CC	-1- SUBCELLULAR LOCATION: COTYLEDONARY MEMBRANE-BOUND VACUOLAR PROTEIN				
CC	BODIES.				
CC	-1- SIMILARITY: NO OTHER 7S SEED STORAGE PROTEINS (PHASEOLIN, VICILIN,				
CC	CONVICILIN, CONGLACININ, ETC.).				
CC					
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/				
CC	or send an email to license@sib-sib.ch).				
CC					
DR	EMBL: M16891; AAA33071.1; -				
DR	PIR: A30838; FMCNAB.				
DR	HSSP: P50477; ICAN.				
DR	INTERPRO: IPR001113; -				
DR	PFAM: PF00546; Seedstore_7s; 1.				
DR	Seed storage protein; Signal.				
FT	SIGNAL	1	25		
FT	CHAIN	1	588		
SO	SEQUENCE	588 AA;	69729 MW;	63699829A	B8ADEB CRC64;

Query Match 32.3%; Score 1074.5; DB 1; Length 588;
Best local similarity 38.2%; Pred. No. 1.2e-59;
Matches 229; Conservative 115; Mismatches 181; Indels 75; Gaps 13;

QY	35	EDPQTECQOCQRRRCQOESDPQOQYQCRCKEICEEENROR--DPQOYEQCQKRC	92
DB	35	DDEPKRYEDCRRCEWDTGQKEQOOCCESSCKSOYGEKQOQOHRREDPQRRRECCQOC	94
QY	93	ORRETEPRHMQICQRCCEERYEKEKKQOKRYEEOQREDEKYEEMKCKDNKRPQOKE	152
DB	95	--RQEEKRPQOCQRCRLRFQEOQO-----SQHQ 123	

OY	153	YEDCRHDEEOD--PRLOVCOORCOEORHGRGDMLNPNOR-----GGSGRVEE	204
OY	154	FOEQCHQHOQORERKQOCYCBREYQE-----NPNRRREDEAELEETEBEQ	175
OY	205	KOSDNPYFDEKSLSTRFTEBGHISVLENFYGRSKLRPAKNRVLLEANPAVFLPT	264
OY	176	BOSHNPFPFHRHSFOSRFEEREGNERVYLRQAPSRHPIRLINEFLSLLEANPNFVLP	235
OY	265	HUDAAILLVGGRCALMIHRDNEESYNLEGGVIRIPAGTFPLIRDNNEHLIAKF	324
OY	236	HCDAKITLVNKRRTLLFLTHENKESINVPGVVVRPAGSTVLANODNKEKLIVL	295
OY	325	LOTSTPGQYKEFPFAGQNPPEVLTSTFKSTLEALNTQTERLGVLG-----QORE	377
OY	296	HRPVNPNQFEFEFFPAGSORPOSYLRAFSRELTLEPAFMTRSQDELDFRGSHRROGQ	355
OY	378	GVIIIRASQOIRRELRDSESRMHIIRRGESRQPYLNFKNRPLYSKKYGQAYEKED	437
OY	356	GNFRASQOIRALSOEATSPR---EKSGE--RFAFMILKTPPYXSNONGFYBACP	409
OY	438	YROLDMOVSVFIANTTQSGMMKPEFNTNSTPVVVVVASGSDVEMACPHLSGRHGRCGG	497
OY	410	FRQSLDINVTASALDONGSIFVPIYNSKATPFVVLVNGNGVENVSHLPQSSFFEEE	469
OY	498	KRHEEEY-----HYQVRRARLSKRAIYVLVLAGPVVFWSSGNELLLFAFG-----I	546
OY	470	EQOQEOEOEEERSGQYKRIQSLSRGDIFVVPANFPYTFVVASQONQNTMTGFLYNO	529
OY	547	NAQNNHNPFLAGRENVLOQIEPQAMELAPASRKEVEELFNGQDESLFFPGRHOOS	606
OY	530	NPDHQRFLVAGKIMV--RQMDSQKELAFGVSSRLVDEIFNNNOESTFYVS--KORQAS	587
RESULT	2		
VCIA_GOSHI			
ID	NC	VCIA_GOSHI	STANDARD; PRT: 605 AA.
NC	P09799;		
DT	01-MAR-1989	(Rel. 10, Created)	
DT	01-MAR-1989	(Rel. 10, Last sequence update)	
DT	15-JUL-1999	(Rel. 38, Last annotation update)	
DE	VICILIN GC72-A	PRECURSOR (ALPHA-GLOBULIN A).	
OS	Gossypium hirsutum	(Upland cotton).	
OC	Eukaryota; Viridipladens; Embryophyta; Tracheophyta; Spermatophyta;		
OC	Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;		
OC	Malvales; Malvaceae; Gossypium.		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RA	Chian C.A., Borroto K., Kamalay J.A., Dure L. III;		
RT	"Developmental biochemistry of cottonseed embryogenesis and		
RT	germination. XIX. Sequences and genomic organization of the alpha		
RL	globulin (vicilin) genes of cottonseed."		
RL	Plant Mol. Biol. 9:533-546(1987).		
CC	-1 FUNCTION: SEED STORAGE PROTEIN.		
CC	-1 SUBCELLULAR LOCATION: CYTOSOL; MEMBRANE-BOUND VACUOLAR PROTEIN		
CC	BODIES.		
CC	-1 SIMILARITY: TO OTHER 7S SEED STORAGE PROTEINS (PHASEOLIN, VICILIN,		
CC	CONVULCIN, CONVULXININ, ETC.)		
CC	-----		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration		
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -		
CC	the European Bioinformatics Institute. There are no restrictions on its		
CC	use by non-profit institutions as long as its content is in no way		
CC	modified and this statement is not removed. Usage by and for commercial		
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/		
CC	or send an email to license@sib-sib.ch).		
CC	-----		
DR	EMBL; M19378; AAA33069.1; -		
DR	PIR; S06398; S06398.		
DR	HSSP; P50477; ICAX.		
DR	INTERPRO; IPR001113; -		
DR	PFAM; PF00546; Seedscore_7s; 1.		
DR	Seed storage protein; Signal.		
FT	SIGNAL		
FT	1		
FT	23		

FT	CHAIN	24	605	VICILIN GC27-A
SO	SEQUENCE	605 AA:	71049 MW:	C9DB9371C976953B CnC64:
	Query Match	32.0%;	Score 1063.5;	DB 1: Length 605;
	Best Local Similarity	38.0%;	Pred. No. 5,9e-59;	
	Matches 223;	Conservative 121;	Mismatches 195;	Indels 63; Gaps
OY	33	NOEDPOTEGCCGCGRCRQGDSDPRQOQYGRKCKELCEEEEEYNORORPOOYECCGCRK	92	
DB	33	SEDDPQREYCECRKRCCLERGTGTEDCKEDSDSEYOLKEEQORDE-DRORYDCRGIC	91	
OY	93	QRRETERP-HMOICQOFCERREYKEXKROOKRYEEQORDEKYEERMKEGDNKRDPOQR	151	
DB	92	QOERRLRLPH---CEGSCRGQYEX-----QQQQQGD-----K	120	
OY	152	EYECRRHCEOE--PRLYOCORCOEDORONGRCGDLMPOR-GSGRVEEGEKOSD	208	
DB	121	QFKCCQRCQMOEDRPRKQOOCVCEKREYOEDPYMGKEENMKEEEEDSDGEOORN	180	
OY	209	NPYEDRSISTRETEEGHIVLENFVGRSKRLALKYURLVLLLEAMPNMFVLTILDA	268	
DB	181	NPYFNHRSQERFREHNGFRLQRFADKHHLLKRIINFRAILEANPNTFVLRHCD	240	
OY	269	DALLVIGRGALKMIHRDNRESYNLECGDVIRTPAGTFYILINDNNEHLIAKFLQTI	328	
DB	241	EKIVVYNGRGTVTFVYHNEKESYNVPGVVYRIPAGSTVYLANODNREKLTIAVLRPV	300	
OY	329	STPOYKEFPFAGGONEPRLTFSKEILLEALNQTETLRGLV-----QOREGVII	381	
DB	301	NNPQCFKFPFAGGONFQSTLRFSHELLLEAFNTRSBOLDPCGROSHRRQSGQGR	360	
OY	382	RASQERLETRDSESRMHIRGGESSRG-PYNLFNRPRLYSKNGOAYEVEKPEDY-R	439	
DB	361	KASQERIALSGATSPR-----KGSBGYAFNLLSPTPSYNOGRFEDCAPRNQ	413	
OY	440	QLODMQVSVFLANTQSSMMGPFPTNTRSTKYVYVVASGADVEMACPHLSGRHGGCGGR	499	
DB	414	QLRVEDSSVAAFELKESIIIVPHYNSKATFVVLVTGEGHVMVCPHLS-ROSSDMSRE	472	
OY	500	HEEEVE-----HYEYVARLSKREALIVLAGHVVNVSGGNNLLLFARGI-NAQNN	551	
DB	473	EEDEDEBERRSQYKRVAAQSLSTGNLFPVAGHGVTFVVASONEDLGLGFLGYNGDN	532	
OY	552	HENFLAGERNVLOQLEPQAMELAPASRKEVEELFNSODESIFPPGPRQHQOQSPSTK	611	
DB	533	KRIYVAKTNNV-RQMDROAKELFGVESRLYDEVFNNNPOESYFVSGHRRGDEDR-RG	590	
OY	612	QOQPLVSLDIF	622	
DB	591	SNMPLSPDLDF	601	
RESULT	3			
GLBL	MAIZE	STANDARD:	PRT:	573 AA.
AC	P15590:			
DT	01-APR-1990 (Rel. 14, Created)			
DT	01-AUG-1990 (Rel. 15, Last sequence update)			
DT	15-JUL-1999 (Rel. 38, Last annotation update)			
DE	GLOBULIN-1 S ALLELE PRECURSOR (GLBI-S) (7S-LIKE)			
GN	GLBI.			
OS	Zea mays (Maize).			
OC	Eukaryota: Viridiplantae: Embryophyta: Tracheophyta: Spermatophyta:			
OC	Magnoliophyta: Liliopsida: Poales: Poaceae: Zea.			
RN	SEQUENCE FROM N.A.			
RP	STRAIN-CV. INBRED LINE VA26:			
RA	Belanger F.C., Kriz A.L.:			
RT	"Molecular characterization of the major maize embryo globulin encoded			
RL	by the Glb1 gene."			
RN	Plant Physiol. 91:636-643(1989).			
	[2]			

OY	377	EGGIIASOEOLREITRDDSESRMHIRPGCGSSSG---- <p>-----PYNLFNKRPLYSNKYGA</p> <p>430</p>
Db	360	MNEGIVAVKSEHYEBELTKAKS-----YSKSGSEEGDITNPINLREGEDLSNNFGKL <p>414</p>
OY	431	YEYKPEDYR-OLQMDVSVFIANTIGSGMMGPFNTSTKVVVVVAGSEADENACPHLSC <p>489</p>
Db	415	FEVPRDKNPLOLOLDMMLTCVEIKEGALMLPHNSKAMVIVVYKGTGMLEIYAVRKEQ <p>474</p>
OY	490	RHGRGGGKRNE-DEEYVHYDOVR---ARLSKREAVYLVAGHPVYVSSGNENILLFAFG <p>545</p>
Db	475	QORGRREDEDEDEDEESNRVRYRTRALKEGVFIIMPAHNPALINASSETHLL--GFG <p>522</p>
OY	546	INAAONHNFNLAGEIRNVLAQOIEPQAMELAFASRKREVELEFNQDSDEIFPFGPRHQOQ <p>605</p>
Db	533	INAEENNHRIFLAGQDNVYIDQIEKQARDLAFPGSGEVEKLIKQKESHFVSARPOQSOQ <p>592</p>
OY	606	SPRSTKQOQ 615
Db	593	SPSSPEKESP 602

RESULT	5	AH11_ARAHY	STANDARD:	PRT:	614 AA.
ID	1	AH11_ARAHY			
AC	1	P43237;			
DT	1	01-NOV-1995 (Rel. 32, Created)			
DT	1	01-NOV-1995 (Rel. 32, Last sequence update)			
DT	1	01-NOV-1997 (Rel. 35, Last annotation update)			
DE	1	ALLERGEN ARA H 1, CLONE P17 (ARA H 1).			
OS	1	Arachis hypogaea (peanut).			
OC	1	Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;			
OC	1	Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I;			
OC	1	Fabales; Fabaceae; Papilionoideae; Arachis.			
RN	1	[1]			
RP	1	SEQUENCE FROM N.A.			
RC	1	STRAIN=CV, FLORUNNER;			
RA	1	MEDLINE=96013631; PubMed=7560062;			
RX	1	Burks A.W., Cockrell G., Stanley J.S., Helm R.M., Bannon G.A.;			
RT	1	"Recombinant peanut allergen Ara h I expression and IGE binding in			
RL	1	patients with peanut hypersensitivity.";			
CC	1	J. Clin. Invest. 96:1715-1721(1995).			
CC	1	-I- SMILARITY: TO 7S SEED STORAGE PROTEINS (PHASEOLIN, VICILIN,			
CC	1	CONVULCIN, CONGLYCININ, ETC.).			
CC	1	-----			
CC	1	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	1	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	1	the European Bioinformatics Institute. There are no restrictions on ways			
CC	1	use by non-profit institutions as long as its content is in no way			
CC	1	modified and this statement is not removed. Usage by and for commercial			
CC	1	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	1	or send an email to license@isb-sib.ch).			
CC	1	-----			
DR	1	EMBL; L38853; AAA60336.1; -.			
DR	1	HSSP; P50477; ICAM.			
DR	1	INTERPRO; IPR001113; -.			
DR	1	PFAM; PF00546; Seedstore_7s; 1.			
DR	1	Allergen.			
QO	1	SEQUENCE. 614 AA; 70283 MW; 1DDACF217ECC5F31 CRC64;			

Query Match	25.98;	Score 861;	DB 1;	Length 614;
Best Local Similarity	34.58;	Pred. No. 2.1e-46;		
Matches 220;	Conservative 114;	Mismatches 205;	Indels 98;	Gaps 23

OY SKYNQDNPQEC-QOCORRROESPPOOXYCQRCKICEEENYNNRPOOOVEB 87
+ + :
29 SPVKTENP---CAQRRLQSC-QDEPDLLKQKACESCTYL-----PRCYVD- 72
Db

OY CQRKCQRREPRIMQICQRCNCRRYEKERRKQOKRYEEQODEEKEFERMKEGD--NK 145
88
73 -----TGATNRHP-----GRTTGRCPGDV-----DDRRQRRREGGRWGP 111
OY

OY	146	KDPOOREVEDCRHRCOEDEPLOTQOCROCOEO-----POHRCGLDLMNFOHRCSGHYEE	201
Db	112	AEPRRREEE-----DMRQPRRDM---RRSIOQPKRIREDGREGE-----QEWGTPRESEV	159
OY	202	GEKOSQNDPYFEDERSLSTRETEEGHISYLENEFYGRSKLLRALKNYRLVLEANPAFY	261
Db	160	REETSRRNPEFFPFRSRESTRGONONGRIYLQREDOQSKOFOMLOHNRIVOIEARPTTLV	219
OY	262	LPTHLDADAILLVYGGGALKMHRDRRESYNLECCGVRVHPAGTPYLLRNONREBLT	321
Db	220	LPKHADADNIIIVIOGGQATVYVANGNKRKSPNLDGALALTPSGFISYILNRIDNOILRY	279
OY	322	AKPLQTIISTPGQYKEFPAGQNPBYLLSTFSKEILEALNTOYERLQVY-----	372
Db	280	AKISMPVMTPOQEFEDFPASSRDOSSYLQSGFSRNTLEAANAEPNEIRVLYLEENAGQO	339
OY	373	---GOOR-----EGYIIASQEOJRELTRDSESRMRHHRGSSS---RQPNYLFNK	419
Db	340	EEORQRRSTRSSNNEGIVAKVSKSEHVOELTKNAKSVS---RKGSEEDITPMLNRDQ	395
OY	420	RPLYSNKGQAYEVKRPEDYR-QLODMOVSYFIANTITGSMAGFFFNRTSRVVVVASGEA	478
Db	396	EPDLSNNNGRLEEVKRPDKKNPQLODDMLMLTCVEIKGKDALRPHNSKANVIYVAKKGTQ	455
OY	479	DYENACPHLSRHRGGRGGKRRHEEEBYNHE---OYR---ARLSKREALVYLGRPYVF	531
Db	456	NLELVAARKEEQQGRREDEWEEEEEEDEEEBGSNREVRATYARKEDVYIMPAHPRVAL	515
OY	532	VSSGNENLELPAFISINQONNENLFLARERVLQOIEPQAMALEFAASRKEVEELEFNFSQ	591
Db	516	NASSRLHL---GFGIMNENNRIRIFLADDKDQVLDQIEKQAKDLAFPOSSGOYERKLLIKNR	573
OY	592	ESIEFPFG-PROHQOQSPRSTQOO-----PLVSYIL	620
Db	574	ESHFVSARPOQSPSPSEKEDEEENQGGKGPLLSIL	610

RESULT	6
ID	GLCA_SOYBN
AC	GLCA_SOYBN
DT	13916,
DT	01-JAN-1990 (Rel. 13, Created)
DT	01-APR-1990 (Rel. 14, Last sequence update)
DT	15-DEC-1998 (Rel. 37, Last annotation update)
DE	BETA-CONGLYCININ, ALPHA CHAIN PRECURSOR.
OS	Glycine max (Soybean).
OC	Eukaryota, Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC	Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I;
OC	Fabales; Fabaceae; Papilionoideae; Glycine.
RP	[1]
RP	SEQUENCE FROM N.A.
RC	TISSUE=COYLEDON;
RX	MEDLINE=91355860; Pubmed=2103438;
RA	Sebastian F.L., Farrel L.B., Schuler M.A., Beachy R.N.;
RT	"Complete sequence of a cDNA of alpha subunit of soybean beta-
RT	conglycinin";
RL	Plant Mol. Biol. 15:197-201(1990).
CC	-1- FUNCTION: SEED STORAGE PROTEIN. ACCUMULATES DURING SEED
CC	DEVELOPMENT AND IS HYDROLYZED AFTER GERMINATION TO PROVIDE A
CC	CARBON AND NITROGEN SOURCE FOR THE DEVELOPING SEEDLING.
CC	-1- SUBUNIT: THE ALPHA-, ALPHA-, AND BETA-SUBUNITS ASSOCIATE IN
CC	VARIOUS COMBINATIONS TO FORM TRIMERIC PROTEINS.
CC	-1- SUBCELLULAR LOCATIONS: EMBRYO AXIS, AND COTYLEDONARY MEMBRANE-BOUND
CC	VACUOLAR PROTEIN BODIES.
CC	-1- SIMILARITY: TO OTHER 7S SEED STORAGE PROTEINS (PHASEOLIN, VICILIN,
CC	CONVICTILIN, CONGLYCININ, ETC.).
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC	the European Bioinformatics Institute. There are no restrictions on its
CC	use by non-profit institutions as long as its content is in no way
CC	modified and this statement is not removed. Usage by and for commercial
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/

[illegible]

Accession	Gene	Protein	Signal	Family
PF00546	Seedstore7s	1	Glycoprotein	Multigene family.
KW	Seed storage protein	Signal	1	
FT	SIGNAL	1	23	BY SIMILARITY.
FT	PROPEP	24	25	POTENTIAL.
FT	CHAIN	26	439	BETA-CONGLYCININ, BETA CHAIN.
FT	CARBOHYD	351	351	N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ	SEQUENCE	439 AA	50552 MW	DBDAA2A077608BB CRC64.
Query Match		22.5%	Score 749	DB 1; Length 439;
Best Local Similarity		37.8%	Pred. No. 1.2e-39;	
Matches 162;	Conservative	87;	Mismatches 146;	Indels 34; Gaps
OY	205	KOSDPYVF-DERSLSTFRTFEEGHISVLENFVGRSKLLRAKRYRLVLEANPNAPVL	263	
DB	28	EDENNPFFFRSSNQTLFENQWNRILRLQFNKRSPQLEMLRYRYQOSKRPNTLLP	87	
OY	264	THLDADAILVIGRGALKMIHRDNRESYNLECGDVIRIPAGTFVYLINDNNERTLIAK	323	
DB	88	HHADDFLLFVLSGRALITLVNNDSDRSYLNHPGDAORIPAGTYVYLVNPHDQNLKIK	147	
OY	324	FLQITSPISGVYKKEPPAGQGNPEYLSFSEKELLEALNMTETRLQV-----GGORE	377	
DB	148	LAIPNKGRIIDDFLSTQAOQSTLQGSFNILLETSHSEFEENINRLVEEEDQOE	207	
OY	378	GVIIIRASQIIRRELRDSESRMRHIRGSGSSRGVYLFNFKRPLYSNKYGQAYEVPED	437	
DB	208	GIVLSEKQKQOEPLVGRRAKSSSKRTI-----SSEDEPNLRSNPIYSNNKGFETREP	263	
OY	438	YRLQDDMVSVFIANTITGSMKGFENFTRSKYVVVAVSGEADVMACPHLSGRHGRGCG	497	
DB	264	NQLRDLDLFLSSVDINGALLPFIKSKALVILVINEGDANIELV-----G	310	
OY	498	-----KRHEEEVHYEQVBARLSKRAIVLAGHPVFEVSSGENMLLFAFCINQNN	551	
DB	311	IKEQQKKQKQOEPLVGRRAKSSSKRTI-----SSEDEPNLRSNPIYSNNKGFETREP	263	
OY	552	HENFLAGERNVLOQIEPQAMELAPASRKKEVEELFNSODESIFPGPQHQQSPSTK	611	
DB	369	ORNFJAGERKDVNVRIEROVELAFPGSADQVERLLKQKRESYFVDA--QPOQKEESKG	426	
OY	612	000PLVSTL 620		
DB	427	RKGFPSTL 435		
RESULT 12				
VL_VICFA				
ID_VCL_VICFA	STANDARD;	PRT;	463 AA.	
AC	P08438:			
DT	01-AUG-1988 (Rel. 08, Created)			
DT	01-AUG-1988 (Rel. 08, Last sequence update)			
DT	01-JAN-1990 (Rel. 13, Last annotation update)			
DE	VICLIN PRECURSOR.			
OS	Vicia faba (Broad bean).			
OC	Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;			
OC	Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I;			
OC	Fabales; Fabaceae; Papilionoideae; Vicia.			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-CV. MINOR.			
RX	MEDLINE=88096511; PubMed=3697075;			
RA	Weschke W., Baumelein H., Mohus U.;			
RT	"Nucleotide sequence of a field bean (Vicia faba L. var. minor) viclin			
RT	gene.";			
RL	Nucleic Acids Res. 15:10065-10065(1987).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-CV. FRIBO;			
RX	MEDLINE=88067789; PubMed=3684610;			
RA	Bassuener R., van Nong H., Jung R., Saalbach G., Muentz K.;			
RT	"The primary structure of the predominant viclin storage protein			
RT	subunit from field bean seeds (Vicia faba L. var. minor cv. Frifo)."?			

RL Nucleic Acids Res. 15:9609-9609(1987).
CC -1- FUNCTION: SEED STORAGE PROTEIN.
CC -1- SUBCELLULAR LOCATION: COTYLEDONARY MEMBRANE-BOUND VACUOLAR PROTEIN BODIES.
CC -1- SIMILARITY: TO OTHER 7S SEED STORAGE PROTEINS (PHASCOLIN, VICILIN, CONVICILIN, CONGLICININ, ETC.).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb.ch/announce/> or send an email to license@isb-sib.ch).
CC -----
DR EMBL: Y00506; CAA68559.1; -
DR EMBL: Y00462; CAA68525.1; -
DR PIR: S06309; S06309.
DR PIR: A27288; A27288.
DR HSSP: P50477; ICAM.
DR INTERPRO: IPR001113; -
DR PFAM: PF00546; Seedstore_7s; 1.
KW Seed storage protein: Signal.
FT SIGNAL 1 27
FT CHAIN 28 463 VICILIN.
FT CONFLICT 257 257 R -> K (IN REF. 2).
FT CONFLICT 443 443 R -> K (IN REF. 2).
SQ SEQUENCE 463 AA; 52694 MW; 5CD09EFDE0D5E6B5 CRC64;

Query Match 21.58; Score 713.5; DB 1; Length 463;
Best Local Similarity 36.0%; Pred. No. 2.1e-37;
Matches 158; Conservative 87; Mismatches 157; Indels 37; Gaps 6;

QY 208 DNPYFEDERSLSTRFTEEGHSVLENFYGRSKLRALKNYRLVLLLEANPAFLPHTD 267
DB 33 DNPVEFSNRPQTLFENENGHTRLOKFDHSLKLENLQNYRLLEYSKPHPTFLPQOTD 92
QY 268 ADAILLVIGRGALKMTHRDRESYNLECGDVIRIPAGTFYLLNDRNNRLHIAFLQT 327
DB 93 ADFLVLVLSGKALITVLPDRNSFSLERGDITKLPAGTIGLYVNRDDEDLVLDLVIP 152
QY 328 ISIPGOYKEFFPAGGONPEPYLSTFSEKLEALINTOTERLRYVQG----- 374
DB 153 VNRGEPOSLELSSGNQNPSTLSGFSKNILEASFNTDYKEIEKVLLEHGKERYHRGLK 212
QY 375 -----GEGVIRASQEQIRELTRDDESRRMHIRGSGSSRGPNLFRKRPYSNKG 428
DB 213 DRROGOEENVIYKISKQJLELNKAKSSS---KSTSESEFPNLRREPLYSNKFG 268
QY 429 QAVEVKEPEDYROLQDMQVSVFIANTIGSMGPFENRSTKVVVVASGEADVEMACPHLS 488
DB 269 KFEFETPKRNPQLODNLIFVNYVEINSGSLLLPHYNSRAIVTYVNGCKDFELVQGRNE 328
QY 489 GRHGGGCGKRHEEEVHYEYR-----ARLSKREAIIVLAGHPVVFSSGENELL 541
DB 329 NOOGLR---EYDEKEKGGEIEIRKQYQAKLSPGDVIVIPAGYFVAITAKSSNLLV- 384
QY 542 FAFGINONHNENFLAGRENVLOQIEPQAMELFAFASRKEVELFNSQDESIFFPQPRQ 601
DB 385 -GGINAENNRFTLAGEEDNVISQIHKPKELAFPGSAOEVDTLLENQKSHFANQPR 443
QY 602 HQOQSPRSTKOQPLVSL 620
DB 444 EREGSGEIKDH--LYSIL 460

RESULT 13
ID VCLB_PEA STANDARD; PRT; 410 AA.
AC P02854;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)

DT 01-JAN-1990 (Rel. 13, Last annotation update)
DE PROVICILIN PRECURSOR (TYPE B) (FRAGMENT).
OS Pisum sativum (Garden pea).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta; OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; OC Fabales; Fabaceae; Papilionoideae; Pisum.
RN [1]
RP SEQUENCE FROM N.A. (CLONES PD087 AND PD084).
RA STRAIN-CV, FELTHAM FIRST;
RX MEDLINE=83220791; Pubmed=6687941;
RA Lycett G.W., Delauney A.J., Gatehouse J.A., Gilroy J., Croy R.R.D., RA Boulter D.;
RT "The vicilin gene family of pea (Pisum sativum L.): a complete cDNA RT coding sequence for preprovicilin.";
RL Nucleic Acids Res. 11:2367-2380(1983).
CC -1- FUNCTION: SEED STORAGE PROTEIN.
CC -1- SUBCELLULAR LOCATION: COTYLEDONARY MEMBRANE-BOUND VACUOLAR PROTEIN BODIES.
CC -----
CC -1- SIMILARITY: TO OTHER 7S SEED STORAGE PROTEINS (PHASCOLIN, VICILIN, CONVICILIN, CONGLICININ, ETC.).
CC PIR: A03344; FWPMBV.
DR HSSP: P50477; ICAM.
DR INTERPRO: IPR001113; -
DR PFAM: PF00546; Seedstore_7s; 1.
KW Seed storage protein; Glycoprotein; Multigene family; Signal.
FT SIGNAL 1 15
FT CHAIN 16 >410 PROVICILIN.
FT SITE 221 222 CLEAVAGE (POTENTIAL).
FT CARBOHYD 359 359 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT NON-TER 410 410
SQ SEQUENCE 410 AA; 46385 MW; 8AF68CE85A316FA2 CRC64;

Query Match 21.4%; Score 712.5; DB 1; Length 410;
Best Local Similarity 37.9%; Pred. No. 2e-37;
Matches 151; Conservative 76; Mismatches 142; Indels 29; Gaps 5;

QY 208 DNPYFEDERSLSTRFTEEGHSVLENFYGRSKLRALKNYRLVLLLEANPAFLPHTD 267
DB 20 ENFIFKSNRNFQTLFENENGHTRLOKFDKRSKIFENLQNYRLLEYSKPHPTFLPQYD 79
QY 268 ADAILLVIGRGALKMTHRDRESYNLECGDVIRIPAGTFYLLNDRNNRLHIAFLQT 327
DB 80 ADFLVLVLSGKALITVLPDRNSFSLERGDITKLPAGTIGLYVNRDDEDLVLDLVIP 139
QY 328 ISIPGOYKEFFPAGGONPEPYLSTFSEKLEALINTOTERLRYVLOOR----- 376
DB 140 VNRGOLQSLFSLGTQKSSLSGFSKNILEAFNTYEIEKVLLEQDEQEPQHRSLK 199
QY 377 -----EGVIRASQEQIRELTRDDESRRMHIRGSGSSRGPNLFRKRPYSNKG 428
DB 200 DRROGINEENVIYKISQJLELNKAKSSS---KSSVSESGPNLRSRNPYSNKG 255
QY 429 QAVEVKEPEDYROLQDMQVSVFIANTIGSMGPFENRSTKVVVVASGEADVEMACPHLS 488
DB 256 KFEFETPKRNPQLODNLIFVNSVDIKVSLLLPHYNSRAIVTYVTEGKDFELVQGR- 313
QY 489 GRHGGGCGKRHEEEVY--HYEQVARLSKREAIIVLAGHPVVFSSGENELLFAFGI 546
DB 314 NENQKENDKEEDEETSQOVOLYRAKLSPGDVIVIPGHVPAIVNNSDNLII--GLGI 371
QY 547 NAONHNENFLAGRENVLOQIEPQAMELFAFASRKEVE 584
DB 372 MAENNERNFLAGEEDNVISQVERVYELAFPGSSHEVD 409

RESULT 14
ID CANA_CANEN STANDARD; PRT; 445 AA.
AC P50477;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)

